

Turing Test, Chinese Room Argument, Symbol Grounding Problem. Meanings in Artificial Agents

Christophe Menant¹

Abstract. The Turing Test (TT), the Chinese Room Argument (CRA), and the Symbol Grounding Problem (SGP) are about the question “can machines think?”. We propose to look at that question through the capability for Artificial Agents (AAs) to generate meaningful information like humans. We present TT, CRA and SGP as being about generation of human-like meanings and analyse the possibility for AAs to generate such meanings. We use for that the existing Meaning Generator System (MGS) where a system submitted to a constraint generates a meaning in order to satisfy its constraint. Such system approach allows comparing meaning generation in animals, humans and AAs. The comparison shows that in order to design AAs capable of generating human-like meanings, we need the possibility to transfer human constraints to AAs. That requirement raises concerns coming from the unknown natures of life and human consciousness which are at the root of human constraints. Corresponding implications for the TT, the CRA and the SGP are highlighted. The usage of the MGS shows that designing AAs capable of thinking and feeling like humans needs an understanding about the natures of life and human mind that we do not have today. Following an evolutionary approach, we propose as a first entry point an investigation about extending life to AAs in order to design AAs carrying a “stay alive” constraint.

Ethical concerns are raised from the relations between human constraints and human values. Continuations are proposed.

1 PRESENTATION

The question “Can machines think?” has been addressed in 1950 by Alan Turing with a proposed test, the Turing Test (TT), where a computer is to answer questions asked by humans. If the answers from the computer are not distinguishable from the ones coming from humans, the computer passes the TT [1]. The validity of the TT has been challenged in 1980 by John Searle with a thought experiment, the Chinese Room Argument (CRA), aimed at showing that a computer cannot understand human language [2]. The possibility for computers to attribute meanings to words or symbols has been formalized by Steven Harnad in 1990 through the Symbol Grounding Problem (SGP) [3].

With the question “can machines think?” understood as “can machines think like human beings think?” [4], we propose to look at these approaches to Artificial Intelligence (AI) by considering that they all address the possibility for Artificial Agents (AAs) to generate meaningful information (meanings) as humans do. The basic question about thinking machines is then

reformulated into “can AAs generate meanings like humans do?” In order to compare meaning generation in humans and in AAs we use an existing system approach to meaning generation: the Meaning Generator System (MGS) where a system submitted to a constraint generates a meaning when it receives information that has a connection with the constraint [5]. We first look at TT and CRA where relations with meaning generation can be considered as rather explicit. The case of SGP is addressed separately as its relations with meaning generation deserve more details. These analysis show that AAs cannot today generate meanings like humans do because human constraints cannot be transferred to AAs. This because we do not understand the natures of life and human mind which are the base ground of human constraints. Consequently, today AAs cannot think as humans do. A better understanding of the natures of life and human mind is needed for designing really intelligent AAs capable of thinking like humans do. We propose an entry point to be investigated: an extension of life to AAs by AAs carrying a “stay alive” constraint.

Ethical concerns are raised as the coverage of human values by human constraints in terms of meaning generation is to be explored.

Continuations are proposed in order to develop with more details several points used here.

2 TURING TEST, CHINESE ROOM ARGUMENT AND MEANING GENERATOR SYSTEM

The TT is about the capability for a computer to understand questions formulated in human language and to answer these questions as well as humans would do. Regarding human language, we can consider that understanding a question is to generate the meaning attached to the question. And answering a question also goes with generating the meaning of the answer. So we can consider that the TT is about meaning generation.

The CRA is to show that a computer not understanding symbols can pass the TT. A person not speaking Chinese and exchanging Chinese symbols with people speaking Chinese can make them believe she speaks Chinese if she chooses the symbols following precise rules written by Chinese speaking persons. The person not speaking Chinese passes the TT. A computer following the same precise rules would also pass the TT. In both cases the meaning of the Chinese symbols is not understood. The TT can be passed without associating any meaning to the exchanged information. The TT does not ensure understanding. Here also, the understanding of the symbols goes with generating the meanings attached to the symbols.

So we can consider that the TT and the CRA are about the possibility for AAs to generate meanings like we humans do.

¹ IBM France (Non Active Employee). christophe.menant@hotmail.fr.

This brings the question about machines capable to think to a question on meaning generation. Can AAs generate meanings as we humans do?

In order to compare the meanings generated by humans and by AAs, we use the Meaning Generator System (MGS) [5]. The MGS models a system submitted to a constraint that generates a meaning when it receives information that has a connection with the constraint. The generated meaning is precisely the connection existing between the received information and the constraint, and it is used to determine an action that will be implemented in order to satisfy the constraint. The MGS is simple. It can model meaning generation in elementary life. A paramecium moving away from acid water can be modelled as a system submitted to a “stay alive” constraint that senses acid water and generates a meaning “presence of acid not compatible with the “stay alive” constraint”. That meaning is used to trigger an action from the paramecium: get away from acid water. It is clear that the paramecium does not possess an information processing system that would allow her to have access to an inner language. But a paramecium has usage of sensors that can participate to a measurement of the acidity of the environment. The information made available with the help of these sensors will be part of the process that will generate the move of the paramecium in the direction of less acid water. So we can say that the paramecium has created a meaning related to the hostility of her environment, in connection with the satisfaction of her vital constraint. Fig 1 illustrates the MGS with this example.

The MGS is a simple tool modelling a system submitted to a constraint. It can be used as a building block for higher level systems (agents) like animals, humans or AAs, assuming we identify clearly enough the constraints corresponding to each case ².

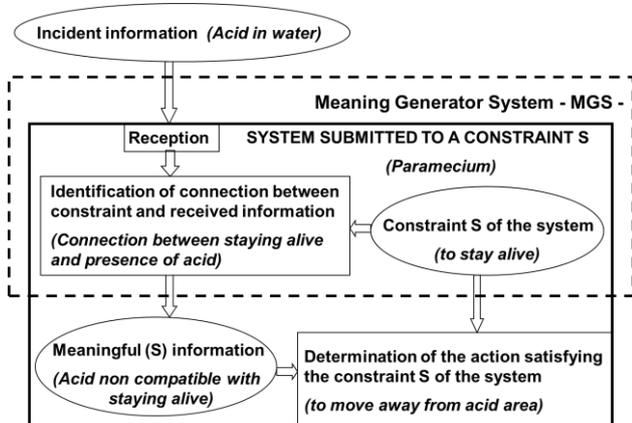


Figure 1. The Meaning Generator System

² The MGS approach is based on meaning generation for constraint satisfaction. It is different from “action oriented meaning”. In the MGS approach, the constraint to be satisfied is the cause of the generated meaning. The action is a consequence of the meaning and comes after it. More on this at [6].

³“Anxiety limitation” has been proposed as a constraint feeding an evolutionary engine in a human evolutionary scenario [14, 15].

The MGS is also usable to position meaning generation in an evolutionary approach. The starting level is basic life with a “stay alive” constraint (for individuals and for species). The sight of a cat generates a meaning within a mouse, as well as a passing by fly within a hungry frog. We however have to keep in mind that staying alive refers to life, the nature of which is unknown as of today. We can easily identify and understand the actions implemented to satisfy a “stay alive” constraint without accessing the nature of the constraint. For humans, the constraints are more difficult to identify as they are linked to human consciousness and free will which are both mysterious entities for today science and philosophy. Some human constraints are however easy to guess like “look for happiness” or “limit anxiety” ³. Reference to the Maslow pyramid can also be used as an approach to human constraints [6]. In all cases, the action implemented to satisfy the constraint will modify the environment, and so the generated meaning. Meanings do not exist by themselves. They come from generation processes that link the agents to their environments in a dynamic mode.

Most of the time agents contain several MGSs related to different sensorimotor systems and different constraints to be satisfied. An item of the environment generates many different interdependent meanings that build up networks of meanings representing the item to the agent. These meaningful representations embed the agent in its environment through constraints satisfaction processes [6].

To see if AAs could generate meanings like humans do, we have to look at how human meaning generation processes could be transferred to AAs. Fig 1 indicates that the constraint is the key element to consider in the MGS. The other elements deal with data processing that is transferrable. When looking at transferring constraints to AAs, we have to recall that the natures of human and animal constraints are unknown as of today. Take for instance the basic “stay alive” constraint that we share with animals. We know the actions that are to be implemented in order to satisfy that constraint, like keep healthy and avoid dangers. But we do not really know what life is. We understand that life came out of matter during evolution, but we do not know how life could be today built up from inanimate matter. We have many definitions for life, but the nature of life is today a mystery. We do not know how to transfer a “stay alive” constraint to AAs. The same applies for human specific constraints which are closely linked to human consciousness. We do not know exactly what is “look for happiness” or “limit anxiety”. We know (more or less) the physical or mental actions that should be implemented in order to satisfy these complex constraints, but we do not know the nature of the constraints. And this is because we do not know the nature of human mind which is, as is the nature of life, a mystery for today science and philosophy. So we have to face the fact that the transfer of human constraints to AAs is not today possible as we cannot transfer things we do not understand.

We cannot today build AAs able to generate meanings as we humans do because we cannot transfer human constraints to AAs. The computer in the TT cannot be today in a position to generate meanings like humans do. The computer cannot understand the questions nor the answers as humans do. It cannot pass the TT. Consequently, the CRA is right. Today AAs cannot think like humans think. Strong AI is not possible today. A better understanding about the nature of life and human mind is

necessary for a progress toward the design of AAs capable of thinking like humans think. Important research activities are in process in these areas [8, 9]. Some possible short cuts may also be investigated, at least for the transfer of animal constraints (see hereunder).

3 SYMBOL GROUNDING PROBLEM AND MEANING GENERATOR SYSTEM

The Symbol Grounding Problem is generally understood as being about how an AA computing with meaningless symbols could generate meanings intrinsic to the AA. "How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?" [3].

This is again about asking how AAs can generate meanings as humans do. The conclusions reached in the previous paragraph apply: AAs cannot today generate meanings as we humans do because we are not in a position to transfer human constraints to AAs. The SGP cannot today have a solution⁴.

Some researchers tend to disagree on the fact that a solution to the SGP can be sufficient for providing meaningful mental states. They consider that meaningful thoughts have to be at the same time conscious mental states [7].

This position can be addressed with the MGS approach where human constraints have to be transferred to AAs so that the AAs can generate meanings like humans do. Such meanings go with human constraints, closely linked to human consciousness.

The requested intrinsic aspect of the semantic interpretation also brings argument in favour of no solution to the SGP when using the MGS approach. Putting aside metaphysical perspectives, we can say that the generation of meaningful information appeared on earth with the first living entities. Life is submitted to an intrinsic and local "stay alive" constraint that exists only where life is. Today AAs are made with material elements that obey physico-chemical laws. The constraints that AAs can carry come from the designer of the AA. The constraints cannot be intrinsic to the AA where there is no life. They are derived from the designer. Only living entities can contain intrinsic constraints and generate intrinsic meanings. Consequently, it is only for living entities that the meaning of the symbols can be "intrinsic to the system". Symbol grounding in a material world does not bring intrinsic meaning generation. AAs can only manage artificial meanings derived from the designer [6]. This comment on the notion of intrinsicness confirms the position expressed above: in the today world of material AAs, the SGP cannot have a solution. We need a better understanding about the nature of life in order to address the possibility for intrinsic meaning generations in AAs. As said, many researches are currently on going on these subjects and there are perspectives for progresses [9].

Another area of investigation for intrinsic constraints in AAs is to look for AAs capable of creating their own constraints.

⁴ Several proposals have been made as solutions to the SGP. Most have been recognized as not providing valid solutions [13].

Whatever the possible paths in this area, it should be clearly highlighted that such approach would not be enough to allow the design of AAs able to think like humans do. The constraints that the AAs might be able to generate by themselves may be different from human ones or managed differently by the AAs. These future AAs may think, but not like humans think. This also brings up ethical concerns for AI where AAs would not be managing constraints and meanings the same way humans do.

4 ARTIFICIAL INTELLIGENCE, ARTIFICIAL LIFE AND MEANING GENERATION

The above usage of the MGS with the TT, the CRA and the SGP has shown that machines cannot today think like humans do because human constraints are not transferrable to AAs.

It is worth recalling that the basic "stay alive" constraint is part of these constraints. And the impossibility to transfer a "stay alive" constraint to AAs means that we cannot today design AAs managing meanings like living entities do. And more basically, not knowing the nature of life makes impossible the design of living AAs. We can only imitate some performances of life. So not only can't we design AAs able to think like humans think, we can't even design AAs able to live like animals live. As shown, the blocking point is our lack of understanding about the natures of life and human consciousness.

In terms of increasing complexity, these subjects can be positioned following an evolutionary approach. It is agreed that life came up on earth before human mind. So it looks logic to address first the problem of the "stay alive" constraint not transferrable to AAs. Even if we do not know the nature of life, we are able to manipulate it. And we could, instead of trying to transfer the performances of life to AAs, look at how it could be possible to extend life with its performances to AAs. Having somehow life living in AAs. Sort of "meat in the computer". The AA would be submitted to the "stay alive" constraints brought in by the living entity, while keeping some control on that living entity. Such approach is different from trying to get organic elements obeying computer logic [10] or trying to use within AAs the sensori-motor performances of living entities, like insect-machine hybrids [11].

The idea of Cyborgs is not new. What we propose to look at is about the possibility to have a living entity resident in an AA and bringing in it a "stay alive" constraint to which the AA would be submitted. This would allow the AA to generate meanings like animals do and interface with its environment like animals do. Such possible progresses in having AAs submitted to resident animal constraints does not bring much about AAs submitted to human constraints. We can however take this as a first step in an evolutionary approach to AAs containing human constraints.

5 MEANING GENERATION, CONSTRAINTS, VALUES AND ETHICAL CONCERNS

The MGS approach has shown that our current lack of understanding about the nature of life and human consciousness makes impossible today the design of AAs able to think or feel like humans do. This because we do not know how to transfer to AAs human constraints that we do not understand. These human constraints are closely related to free will and consciousness that are mysteries for today science or philosophy. But human

constraints do not a priori include human values. Some find happiness with the suffering of others. The way by which human constraints can take into account human values is not obvious, nor clear. This brings to highlight here ethical concerns that address two directions at least.

First, researches about the nature of human consciousness should consider how human values could be linked to human constraints. This is a challenging subject as human values may not be universal. But the nature of human consciousness is still to be discovered and we can hope that its understanding will shed some light on the diversity of human values.

As addressed above, another case is the one about AAs becoming capable of generating by themselves their own constraints. Such approach should be careful keeping human values in the background of these constraints so the AAs are not brought in a position to generate meanings and actions too distant from human values.

6 CONCLUSIONS

Meaning generation for constraint satisfaction shows that we cannot today design AAs capable of generating meanings like humans do. Our lack of understanding about the natures of life and human mind makes impossible the transfer of human constraints to AAs. The consequence is that human type meaning generation is not today possible within AAs. Considering the TT, the CRA and the SGP as being about meaning generation, we can say that today AAs cannot think like we humans do, they cannot pass the TT. The CRA is correct, and the SGP cannot have a solution. Strong AI is not possible. Only weak AI is possible. Imitation performances can be almost perfect and make us believe that AAs generates human-like meanings, but there is no such meaning generation as AAs do not carry human constraints. AAs do not think like we do and have no feeling about what is going on as they do not carry human constraint and cannot generate meanings like we do. Another consequence is that it is not possible today to design living machines, as we do not know the nature of life. True AL is not possible today.

Understandings about the nature of life and human consciousness are needed to design AAs capable of behaving like animals and thinking like humans. As life is less complex and easier to understand than consciousness, the transfer of a "stay alive" constraint should be addressed first. An option could be to extend life with its "stay alive" constraint to AAs. The AA would then be submitted to the constraints brought in by the living entity.

7 CONTINUATIONS

The MGS approach applied to TT, CRA and SGP has highlighted that the constraints to be satisfied are at the core of a meaning generation process. We feel that an evolutionary approach to the nature of constraints would allow an interesting perspective by identifying their origins and their positioning relatively to physico-chemical laws, and so introduce an evolutionary theory of meaning. Work is in process on these subjects [6, 12]. But human constraints remain ill-defined and not really understood. As said, they are tightly linked to self-consciousness and free will which are mysteries for today

science and philosophy. Significant work is to be done in this area, where a better understanding of human mind is needed.

The MGS approach offers the possibility to define meaningful representations that embed agents in their environments [6]. Such representations can be used as tools in an evolutionary approach to self-consciousness where the human constraints play a key role. Work is in process in this area also [12].

An evolutionary approach to human constraints brings to focus first on the ones shared with animals, and more precisely on the "stay alive" constraint. As introduced above, we feel it could be interesting to look at extending life with its performances to AAs in order to bring the "stay alive" constraint to AAs. Some living entity in the AA, with the AA keeping control on it and being submitted to the "stay alive" constraint, like animals are. Investigating such an approach calls for developments which are beyond the scope of this paper.

Ethical concerns have been raised through the possible relations between human constraints and human values. If AAs can someday carry human constraints, they may not carry human values. An evolutionary approach to human consciousness could bring some openings on that perspective. Such concern applies also to the possibility of AAs creating their own constraints that may be different from human ones and consequently not linked to human values.

REFERENCES

- [1] Turing, A.M. (1950). "Computing machinery and intelligence. *Mind*, 59, 433-460.
- [2] Searle, J. R. (1980) "Minds, brains and programs". *Behavioral and Brain Sciences* 3: 417-424.
- [3] Harnad, S. (1990). "The Symbol Grounding Problem" *Physica*: 335-246.
- [4] Harnad, S. (2008) "The Annotation Game: On Turing (1950) on Computing, Machinery, and Intelligence" in: *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, Springer. ISBN 13: 978-1-4020-6708-2
- [5] Menant, C. "Information and Meaning" *Entropy* 2003, 5, 193-204, <http://cogprints.org/3694/1/e5020193.pdf>
- [6] Menant, C. (2010). "Computation on Information, Meaning and Representations. An Evolutionary Approach" in *Information and Computation: Essays on Scientific and Philosophical Understanding of Foundations of Information and Computation*, G. Dodig-Crnkovic, M.Burgin. ISBN-10: 9814295477
- [7] Rodriguez, D. and all: "Meaning in Artificial Agents: The Symbol Grounding Problem Revisited". *Minds & Machines*, Dec 2011.
- [8] Philpapers "Nature of Consciousness" <http://philpapers.org/s/nature%20of%20consciousness>
- [9] Philpapers "Nature of life" <http://philpapers.org/s/nature%20of%20life>
- [10] "Scientists create computing building blocks from bacteria and DNA" *News and Events. Imperial College London. 18 Oct 2011.* http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/newssummary/news_18-10-2011-16-7-29
- [11] "Hybrid Insect Micro Electromechanical Systems (HI-MEMS)". http://www.darpa.mil/Our_Work/MTO/Programs/Hybrid_Insect_Micro_Electromechanical_Systems_%28HI-MEMS%29.aspx.
- [12] Menant, C. "Evolutionary Advantages of Intersubjectivity and Self-Consciousness through Improvements of Action Programs". TSC 2010 <http://cogprints.org/6831/>
- [13] Taddeo, M., Floridi, L. "Solving the Symbol Grounding Problem: a Critical Review of Fifteen Years of Research" *Journal of Experimental & Theoretical Artificial Intelligence*, Volume 17, Number 4, Number 4/December 2005.

- [14] Menant, C. " Information and Meaning in Life, Humans and Robots". *In Procs. Foundation of Information Sciences 2005, Paris*.
<http://cogprints.org/4531/1/F.45.paper.pdf>
- [15] Menant, C. " Evolution and Mirror Neurons. An introduction to the Nature of Self-Consciousness". *TSC 2005*.
<http://cogprints.org/4533/1/Charts.pdf>